# Machine Learning Approaches for Crop Leaf Disease Detection and Classification

Synopsis

Swami Ramanand Teerth Marathwada University, Nanded

FOR THE AWARD OF THE DEGREE OF

**DOCTOR OF PHILOSOPHY**

**IN**

**COMPUTER SCIENCE**

**FACULTY OF SCIENCE AND TECHNOLOGY**

By

**Mr. Mahesh Shyamsunder Darak**

Under the guidance of

**Dr. N. K. Deshmukh**
Assistant Professor
School of Computational Sciences,
S. R. T. M. University, Nanded.

# Table of Contents


| Sr. No | Contents | Page No. |
|---|---|---|
| 1 | Abstract | 1 |
| 2 | Introduction | 2 |
| 3 | Review of Literature | 5 |
| 4 | Objectives of Research | 7 |
| 5 | Methodologies to be adopted | 8 |
| 6 | Importance of Study | 9 |
| 7 | Proposed Work Plan | 10 |
| 8 | References | 11 |

## 1. Abstract:

Crop disease diagnosis is very crucial task for every farmer and individual in order to prevent various losses like less productivity, less quality and quantity or it can also lead to defective yield. Therefore early identification and early detection can help to save the crop yield and also the quantity and quality of the production can be preserved by proper treatment and solution required from time to time. The automation system will help in identification and also diagnose the disease of the crop since crop disease is one of the important factors which can effect on the quantity and quality of the product.

Constant observation of crop from seed germination to the growing of the crop with the automation system will help the farmers to identify and diagnose it very fast as compared to traditional approach of checking every crop leaf and then predicting the disease with his prior knowledge and experience in which there are chances of low accuracy and efficiency of prediction. So to overcome the problems occurred in traditional approach.

Machine learning and image processing can be used to detect the disease and identify the disease on the crop.

## 2. Introduction :

Agriculture is the backbone of Indian economy. In order to survive in such a changing environment and the climatic conditions automation and up- gradation of the traditional approaches and methods of disease detection is essential. The agriculture sector requires extreme up-gradation in order to meet the challenges of demand and supply of particular crop. Crop optimum
yield can be obtained by periodic monitoring of the crop and timely disease identification if infected by any disease.

Crop health is an important factor; if plant is having disease then it is having direct influence to continuous reduction of quality as well as quantity of yield of crop.

In market there are various numbers of pesticides available but selection of the correct pesticide can be done only if proper identification and detection of disease is done then only the pesticide use can be effective to control the effects of the disease on the crop. Thus in order to detect crop diseases accurately ,image processing and machine learning techniques are very much useful and deep learning can also be used for more accuracy and efficient prediction.

### 2.1 Crop Diseases

### a. Cotton Diseases:

1) Alternaria : It is a fungal disease mainly caused by A. Alterneta or Alternaria macrospora [1]. The disease is most severe on the lower part of leaves as compared to the upper part and may get confused with the spots of bacterial leaf blight as the symptoms are nearly similar [2]. At the beginning, brown, gray-brown to tan colored small circular spots appears on leaves and vary from 1-10mm in size which later on become dry, dead with gray centers which crack and fall out [2]. Sometimes, old spots combine together and create irregular dead areas.

2) Cercospora: Cercospora is brought about by the Cercospora Gossypina [1]. The tainted leaf has red spot blemishes on the leaves which expand in distance across to around 2 cm. The spots are round or unpredictable in shape with yellowish, purple, dark brown or blackish fringes with white focuses [2]. The rakish leaf spot shows up because of the limitation of the lesion region by fine veins of the leaf. This malady influences more seasoned leaves of develop plants. 4) Grey Mildew It is fungal disease generally caused by the Ramularia Areola Atk. in imperfect stage and Mycosphaerella areola in perfect stage [1]. At the initial stage, the infection appears as triangular,

square or irregularly circular whitish spots of 3–4 mm size on leaves [2]. This disease primarily appears as irregular angular and pale translucent spots of diameter 1- 10mm on older leaves of matured plant [3]. The infected regions are light to yellowish green in color on the upper surface. As disease infection increases, the small spots merge together forming bigger spots and the leaf tissues turn yellowish brown while whitish frosty growth appears on a lower surface but occasionally on the upper surface [3] and many more diseases like Bacterial Blight, Grey Mildew and Fusarium Wilt.

### b. Wheat Diseases :

Wheat plant Diseases The diseases which affect Wheat plant are (1) Leaf Rust (2) Powdery Mildew (3) Stem Rust and (4) Yellow Rust [4]. Leaf Rust causes small brown lesions on the leaves. These blister like lesions are common on leaves but can occur on the leaf sheath, which extends from the base of the leaf blade to the stem node. Lesions caused by leaf rust are normally smaller, more round, and cause less tearing of the leaf tissue than those caused by stem rust. Powdery mildew causes white lesions on leaves and leaf sheaths. Glumes and awns also can be infected when the disease is severe. Fungal growth is largely limited to outer plant surfaces and can be easily wiped away by rubbing a finger across affected areas. Mature lesions may have dark, reproductive structures mixed with the white, cottony growth of the fungus [5].

### c. Soyabeen Diseases:

Brown Spot (Septoria glycines) :- Brown spot is a common disease which occurs frequently on young plants. If warm moist weather conditions continue, infection, and defoliation progress from lower to upper leaves. Usually the hot dry weather of midseason arrests disease development, but the disease may become active again near maturity. If conditions favorable for disease development continue into the season or reoccur before maturity, the resulting defoliation can cause serious yield reductions. The pathogen overwinters on infected leaf and stem debris, and the disease is most severe when soybeans are grown continuously in the same field. Cotyledons, primary leaves, and lower trifoliolate leaves typically show brown spot symptoms. The brown to red lesions vary in size from pinpoint to 4mm wide, but may coalesce and form larger irregular shaped spots [6].

### d. Tomato Diseases:

1) Leaf mold: The symptoms of this disease show to be different in both sides of the leaf. While yellow spots appear in the front side at early infection, brown sporulation occurs in the back side even ate early stage. It spreads rapidly from plant to plant, covering leaves with lesions and defoliating plants which severely

reduces yield . Leaves tend to roll and dry in the last stage of life. In this disease the fruit is not infected.

2) **Gray mold Foliar**: symptoms are more common under humid greenhouse conditions. Affected leaves show light or gray spots, and the infected areas become covered by a brown fungal growth. This disease occurs in warm and sunny weather and hig humidity conditions, affecting the fruit. One specific characteristic in the last stage of the leave is a white hair on the surface [7].

## 2.2 Machine Learning Methods:

Machine learning, by its definition, is a field of computer science that evolved from studying pattern recognition and computational learning theory in artificial intelligence. It is the learning and building of algorithms that can learn from and make predictions on data sets. These procedures operate by construction of a model from example inputs in order to make data-driven predictions or choices rather than following firm static program instructions. “A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E.” -- Tom Mitchell, Carnegie Mellon University [8].

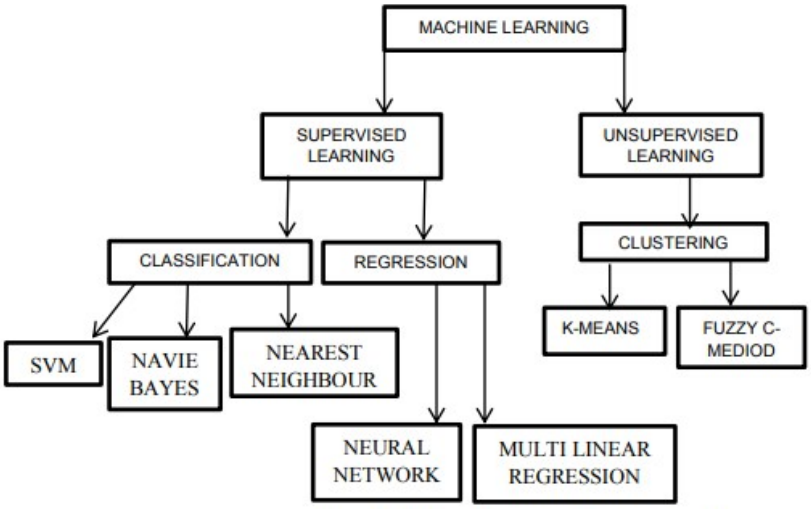



Fig: Block Diagram of Machine learning methods

Machine Learning: Algorithms Types Machine learning algorithms are organized into taxonomy, based on the desired outcome of the algorithm.

Common algorithm types include:

• **Supervised learning**: where the algorithm generates a function that maps inputs to desired outputs. One standard formulation of the supervised learning task is the classification problem: the learner is required to learn (to approximate the behavior

of) a function which maps a vector into one of several classes by looking at several input-output examples of the function.

• **Unsupervised learning**: which models a set of inputs: labeled examples are not available.

• **Semi-supervised learning**: which combines both labeled and unlabeled examples to generate an appropriate function or classifier.

• **Reinforcement learning**: where the algorithm learns a policy of how to act given an observation of the world. Every action has some impact in the environment, and the environment provides feedback that guides the learning algorithm.

• **Transduction**: similar to supervised learning, but does not explicitly construct a function: instead, tries to predict new outputs based on training inputs, training outputs, and new inputs. • Learning to learn --- where the algorithm learns its own inductive bias based on previous experience [9].

## 3. Review of Literature:

A weather based prediction model is proposed by Rakesh et.al.[10]. It is developed using support vector machines for prediction of rice blast disease. Six significant weather variables such as temperature, humidity were selected as predictor variables. The performance of conventional multiple regression, artificial neural network and support vector machine (SVM) was compared. Authors found that SVM based regression approach provides an enhanced description of the association between the environmental conditions and disease level. Authors developed a SVM based web server for rice blast prediction to help decision making process. This suggests that such online systems for all crops will greatly help farmer to improve the productivity.
Remedial measures for pest or disease management using back propagation neural network for recognition of leaves are suggested in [11]. Training was performed using only 5 specimen of each type of disease. Prewitt edge detection and thinning algorithm is used to find leaf tokens. These tokens are fed to back propagation algorithm. Authors reported that there is a space for improvement of this work by more experimentation's with bier train in sets Authors suggested that it is possible to develop an expert system from agricultural research about estimation of damage due to pests and diseases on the crops and remedial measures for it is available.

Camargo and Smith developed method for recognizing the leaf regions affected by disease lesion [12]. Leaves of maize, cotton, soybean banana are used for experimentation. Color image of diseased leaf is transformed into HSV and I1I2I3 color space. For segmentation of disease by intensity thresholding method only H and I3 components of image are used. Authors concluded that developed method

efficiently separates healthy and diseased areas for large variety of conditions and plants.

Viraj Gulhane and Gurjar developed method for detection of diseases on Cotton Leaves [13]. Features of diseased leaf are extracted from color of the disease using modified self-organization feature map and back propagation neural network. SVM is used to classify leaf diseases. Disease detection efficiency of about 86% is achieved by this method.

Dheeb Al Bashish et.al. designed an image processing based system for automatic detection and classification of plant leaf diseases [14]. Five diseases named early scorch, cottony mold, ashen mold, late scorch and tiny whiteness were used for testing the applicability of the system. The leaf images of above diseases are collected and transformed in to device independent color space. Then the leaf images are clustered into diseased and healthy portions using K-means clustering technique. Textures features of segmented diseased portion of leaf are extracted using color cooccurrence matrix in HSV color space. These features applied to pre-trained neural network which automatically detect and recognize the leaf disease. Authors reported that classification precision around 93% is provided by the develop system. Authors also suggested that there is scope for developing better segmentation and feature extraction techniques.

P. Revathi and M. Hemalatha developed method for detecting Cotton leaf spot diseases using Skew Divergence Method [15]. Diseased cotton leaves are preprocessed and diseased regions are segmented. To extract features Enhance Particle Swarm Optimazation based Feature Extraction Method (EPSO) is used. Skew divergence method and features like edge, color, texture variances of the diseased portion of leaf image is used. These features are used to test performance of classifiers such as SVM, Back propagation neural network (BPNN), Fuzzy with Edge CYMK color feature and GA. Authors found that when EPSO algorithm is combined with fuzzy classifiers higher accuracy is achieved

A system for recognition of Cotton leaf diseases is developed using SVMClassifier [16]. Cotton leaf diseases such as Bacterial Blight, Myrothecium and Alternaria have been diagnosed. The diseased leaf images are preprocessed using unsharp filter to enhance edges and se mented usin Otsu’s segmentation. Various color and shape features are extracted and applied as input to SVM classifier which provided average accuracy of 90%. Authors suggested that this system could be developed for detection of other cotton leaf diseases as well as for development of hybrid algorithms such as genetic algorithm and neural network so as to enhance classification accuracy.

Chenghai yang proposed a system describes the design and testing of an airborne multispectral digital imaging system for remote sensing applications. The system consists of four high resolution charge coupled device (CCD) digital cameras [17] this system involved huge maintenance cost.

P. Revathi, et al.[18] proposed two phases to identify the lesion region of the disease. First Edge detection technique is used for segmentation and then image analysis and classification of diseases is done using the proposed HPCCDD Algorithm.

## 4. Objectives of Research:

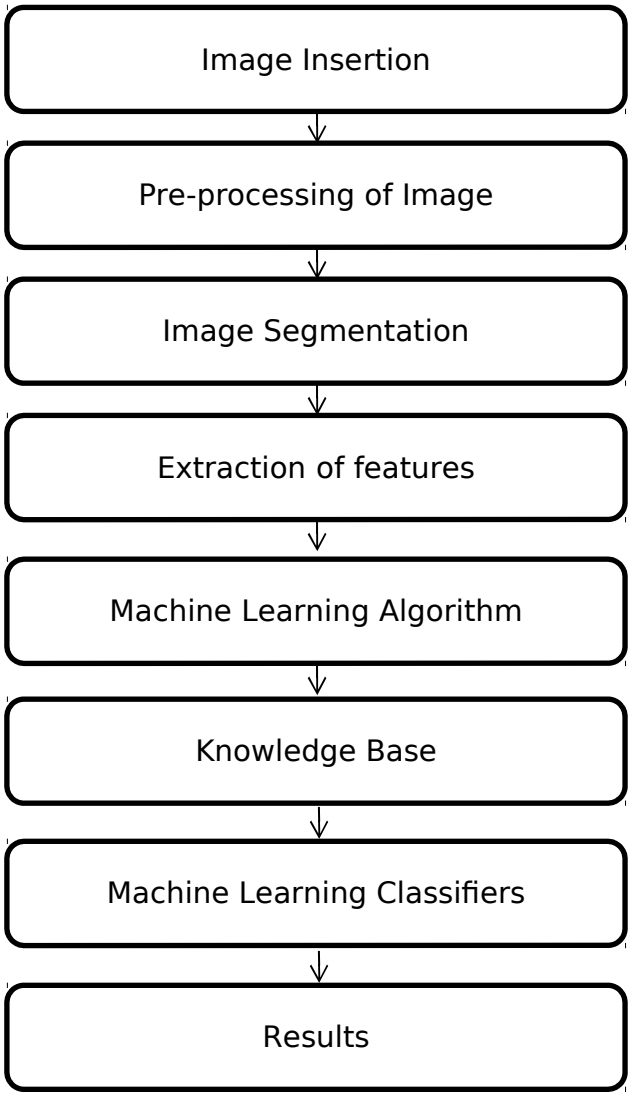
- Image processing has played an important role in the field of Agriculture for estimation of crop yield, vegetative growth of crop and disease detection etc. This research is carried out to study effectiveness of image processing and crop disease detection and identification using machine learning.

- According to Prof. Dr. David who is Guest Editor of 'Impact of Plant Disease on Food Security (2015)', farmers expend billions of dollars on disease management lacking enough technical support. It results in improper disease control, environmental pollution and destructive results of crops. Disease management includes detection of disease, its type, severity and use of pesticides for curing it. There are environmental issues on using pesticides because pesticides may be toxic to some kind of birds, and damage food chains. Hence technical support in this field is essential for early detection of disease and its severity. A usual practice for agricultural experts to detect disease is based on ability to recognize disease and estimate the damage by observation. The damage is estimated without photo documentation. The drawback of this system is subjectivity and low throughput. Electronics or Computer based expert systems can overcome these limitations.

- Electronic systems assist farmers in recognizing or detecting the diseases. These systems help farmers for right decision of proper treatment. These are intelligent computer programs that are competent of contributing solutions to specific problems in given field, in a way and at a level equivalent to that of human expert in that field. These systems work with reduced information compared to human experts. Another benefit of expert system is that it performs tasks more consistently than human experts which increases throughput and reduces subjectivity.

- To be familiar with the disease found in crops
- To be known with different signs and symptoms of the diseases
- To get maximum accuracy to regarding the control measures and management of the disease found in crops.

## 5. Methodologies to be adopted



The proposed system has three phases; first phase deals with image processing using image segmentation extraction and then comes machine learning role. Second phase deals with training datasets. Both healthy and diseased leaf images are collected. Once the dataset is ready with healthy and infected image samples, the threshold is extracted for both aging and for diseases. Periodically images are obtained by remote sensing or dataset images. RGB values of the monitored images are extracted and compared with threshold images. If the threshold is greater or less than given value, histogram analysis and edge detection techniques are used to identify particular plant diseases. Third phase deals with testing using machine learning approaches.

## 6. Importance of Study:

- International and National Status:

  In traditional farming and disease detection the farmer uses his prior experience and knowledge which in turn many times leads to inaccuracy of disease detection and cannot discover correct diseases and measures to control the diseases. Continuous monitoring in farming leads to increase in the productivity of the crop and early detection of disease will help to take precautionary measure on time in order to protect the crop from sever diseases.

  The proposed system will help farmers to periodically monitor the cultivated crop. Machine learning techniques are used to train the model which helps to take proper decision regarding the diseases. The pesticide as a remedy can be suggested to the farmer for the infected disease of the crop to control it.

- Significance of the Study:

  It provides proper disease detection on the crop. It detects and identifies the disease so that proper pesticides can be used as per the disease. The overall goal of the system is to detect the disease and identify which disease it is for proper selection of control measures with respect to the disease occurred which will lead to good quality of crop production and quantity.

## 7. Proposed work plan

- **Year-wise plan of work and targets to be achieved:**

| Task | Year 1 | | | | Year 2 | | | | Year 3 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Q1 | Q2 | Q3 | Q4 | Q5 | Q6 | Q7 | Q8 | Q9 | Q10 | Q11 | Q12 |
| Literature survey | | | | | | | | | | | | |
| Study of Crops | | | | | | | | | | | | |
| Study of Crops & Diseases | | | | | | | | | | | | |
| Study of influence of diseases on the crop | | | | | | | | | | | | |
| Collection of crop images | | | | | | | | | | | | |
| Study of research papers and algorithms used in this area | | | | | | | | | | | | |
| Study of database for crops | | | | | | | | | | | | |
| Finalizing the architecture | | | | | | | | | | | | |
| Extract features from image using techniques | | | | | | | | | | | | |
| Analysing features for verification using methods | | | | | | | | | | | | |
| Derived an algorithm/ methodology for crop disease detection and | | | | | | | | | | | | |

| | | | |
|---|---|---|---|
| identification | | | |
| Documentation, thesis writing, etc. | | | |

## 8. References


[1] Texa plant disease handbook, [online] Available: http://plantdiseasehandbook.tamu.edu/industryspecial tylfiber-oil-specialty/cotton/.

[2] Plant village Cotton, [online] Available: https://www.plantvillage.org/en/topics/cotton.

[3] P. Revathi, M. Hemalatha, "Advance Computing Enrichment Evaluation of Cotton Leaf Spot Disease Detection Using Image Edge detection", IEEE, July 2012.

[4] Huang et al., "New Optimized Spectral Indices For Winter Wheat Diseases", IEEE journal of selected topics in applied earth observations and remote sensing, pp. 128-135, 2014

[5] **Delwiche**, M. S. Kim, "Hyperspectral imaging for detection of scab in wheat", Biol. Qual. Precis. Agric., vol. 4203, pp. 13-20, 2000.

[6] Pratibha Dongre, Mr. Toran Verma , ”A Survey of Identification of Soybean Crop Diseases”, International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 1, Issue 8, October 2012, ISSN: 2278 – 1323.

[7] Alvaro Fuentes and 5 others, “Characteristics of Tomato Plant Diseases”, Conference Paper · October 2016.

[8] Annina Simon, Mahima Singh Deo, S. Venkatesan, D.R. Ramesh Babu,” An Overview of Machine Learning and its Applications”, International Journal of Electrical Sciences & Engineering (IJESE), Volume 1, Issue 1; 2015 pp. 22-24.

[9] Taiwo Oladipupo Ayodele, “ Types of Machine Learning Algorithms”, https://www.researchgate.net/publication/ 221907660_Types_of_Machine_Learning_Algorithms, · February 2010

[10] Rakesh Kaundal, Amar S Kapoor and Gajendra PS Raghava (2006),“Machine learning techniques in disease forecasting: a case study on rice blast prediction”, BMC Bioinformatics, 1-16.

[11] M. S. Prasad Babu, B. Srinivasa Rao (2007), “Leaves Recognition Using Back Propagation Neural Network-Advice For Pest and Disease Control On Crops”, IndiaKisan Net: Expert Advisory System

[12] Camargo A, Smith J.S. (2 ), “An image-processing based algorithm to automatically identify plant disease visual symptoms”, Biosyst En ,1 2, –

[13]. Viraj A. Gulhane, Ajay Gurjar (2011), “Detection of diseases in cotton leaves and its possible dia nosis”, International Journal of Ima e Processin , vol.5, issue 5, 590-598.

[14] Dheeb Al Bashish Malik Braik, Sulieman Bani- Ahmad (2011), “Detection and Classification of Leaf Diseases using K-means- Based Segmentation and Neural Networks based Classification”, Information Technology Journal 10(2), 267-275.

[15]P. Revathi, M. Hemalatha (2014), “Cotton Leaf Spot Diseases Detection Utilizin Feature Selection with Skew Diver ence Method”, International Journal of Scientific Engineering and Technology ISSN : 2277-1581,Volume No.3, Issue No.1, 22- 30. 220

[16] P.R. Rothe, R. V. Kshirsagar (2014), “SVM-based Classifier System for Recognition of Cotton Leaf Diseases”, International Association of Scientific Innovation and Research (IASIR) ,International Journal of Emerging Technologies in Computational and Applied Sciences , ISSN (Print): 2279- 0047 ISSN (Online): 2279-0055 , 427-432.

[17] Chenghai Yang, "A high-resolution airborne four-camera imaging system for agricultural remote sensing", Computers and Electronics in Agriculture, pp. 13-24, 2012.

[18] P. Revathi, M. Hemalatha, "Classification Of Cotton Leaf Spot Disease Using Image Processing Edge Detection Technique", International Conference on Emerging Trends in Science Engineering and Technology IEEE, pp. 169-173, 2012.

| | |
|---|---|
| Signature | Signature |
| Research Scholar | Research Guide |
| Mr. Mahesh S. Darak | Dr. Nilesh. K. Deshmukh |